

Tax Compliance in the New Millennium: Understanding the variables

Sue Yong^{* a}, Karen Lo^{**}, Brett Freudenberg^{***a}, and Adrian Sawyer^{****}

Abstract

Taxpayer compliance is a critical part of any tax system, as without it, the integrity of the tax system can be undermined. Previously, Jackson and Milliron published a comprehensive study reviewing and synthesising the findings of tax compliance studies conducted prior to 1986; highlighting the variables, methodologies and theories. Richardson and Sawyer built upon this research by synthesising the work from 1986 to 1997. Similarly this current research seeks to build upon these two important studies and extend our knowledge by analysing tax compliance studies in terms of ‘variables’ identified from January 1998 to December 2017.

This analysis was facilitated by using text mining software, Leximancer, to visualise the frequency and co-occurrence of concepts and themes related to tax compliance factors. Data was drawn from 713 articles obtained from 15 research databases. Given the size of the data, this article is the first of a series of articles, focussing on the ‘variables’ that have been researched over the 20 year period. Future articles will examine the methodologies and theories used in the literature over this 20 year period.

Keywords: tax compliance, evasion, morale, compliance cost, Leximancer

* AUT University (email: sue.yong@aut.ac.nz);

** AUT University (email: Karen.lo@aut.ac.nz);

*** Griffith University (email: b.freudenberg@griffith.edu.au);

**** University of Canterbury (email: adrian.sawyer@canterbury.ac.nz)

We wish to thank the two anonymous referees that provided valuable comments and suggestions which have significantly improved our article. Any remaining errors are ours.

This article was accepted for publication on 16 October 2019.

Section One

INTRODUCTION

Taxpayer compliance is a critical part of any tax system, as without it, the integrity of the tax system can be undermined, with additional resources expended by revenue authorities and insufficient tax revenue raised. Governments around the world are confronting the problem of raising sufficient tax revenue to fund their public goods and services programs. Politically there can be resistance to simply raising tax rates, and governments are becoming more acutely aware how there can be tax leakage, given the perceived non-compliance by taxpayers. Taxpayers of concern may range from multinational enterprises (MNEs) utilising transfer pricing strategies through to small micro-businesses participating in the cash economy. In 2004 it was estimated that the cash economy can range between 3 to 15% of gross domestic product.¹ A World Bank study of the ‘shadow’ economy in 2007 resulted in estimates of: the United States of America (USA) 8.4%; Japan 10.3%; New Zealand 12.0%; China 12.1%; United Kingdom 12.2%, Australia 13.5% and Canada 15.3%.² If these estimates are accurate, the amount of tax revenue potentially being lost is substantial. Additionally, more governments, especially developed nations, are relying on taxpayers to self-assess their tax liability, and for such systems to be able to operate they require a sufficient level of compliance by taxpayers. Accordingly, tax compliance is an important concept to consider, and is part of the reason that it has been researched through a range of disciplines from psychology to economics and accounting.

¹ (The) Senate (Economics Reference Committee). “The Structure and Distributive Effects of the Australian Taxation System” (Commonwealth of Australia, Canberra, 2004), at 62.

² F Schneider, A Buehn and C.E. Montenegro, “Shadow Economies All over the World New Estimates for 162 Countries from 1999 to 2007”, Policy Research Working Paper 5356, (The World Bank Development Research Group Poverty and Inequality Team & Europe and Central Asia Region Human Development Economics Unit July 2010). However, it should be acknowledged that estimating the cash economy can be problematic, as in 2017, the Black Economy Taskforce in Australia suggested that the Australian cash economy in 2015-2016 may have only been 3% of GDP: see Black Economy Taskforce: Final Report—October 2017 at p 35, as cited in https://www.aph.gov.au/About_Parliament/Parliamentary_Departments/Parliamentary_Library/pubs/rp/BudgetReview201819/TargetingBlackEconomy.

There have been a number of theories postulated to try to explain why taxpayers may or may not comply, extending from deterrence theory³ to motivational postures.⁴ It is likely that no one theory can completely explain the complex relationship between taxpayers and the tax system. There are many types of taxpayers, ranging from salary and wage earners, to high wealth individuals and those conducting enterprises of various sizes, as well as taxpayers with different cultural and religious backgrounds, levels of literacy, and age. Also, the tax culture of a jurisdiction may influence tax compliance, with the presence or absence of trust and/or corruption potentially influencing compliance behaviour.⁵ Furthermore, it could be that these factors alter with time, as there are changes in technologies, working habits and our understanding of tax compliance. Since the 1960s there have been many studies around the world exploring tax compliance in various jurisdictions. In isolation these prior studies have been important in illuminating various factors about why people do or do not comply with the tax system. However, what observations are available if we consider these prior studies as a whole? It is argued in this article that such aggregated reviews of prior studies are important in illuminating important trends and commonalities.

To gain an understanding of our knowledge of tax compliance, Jackson and Milliron published a comprehensive study reviewing and synthesising the findings of mainly USA tax compliance studies conducted in the 1970s to the end of 1985.⁶ In this review, they identified compliance variables, the methodological issues and the theoretical areas considered. This initial study was then built upon by Richardson and Sawyer who analysed subsequent tax compliance research from 1986 to the end of 1997. Richardson and Sawyer sought to determine the extent to which earlier identified compliance variables, methodological issues and theoretical areas continued to be found in studies from 1986 to 1997.⁷ Through this study, Richardson and Sawyer not

³ N Walker, *Crime and Punishment in Britain* (Edinburgh: Edinburgh University Press, 1965).

⁴ V Braithwaite, "Tax and democracy: Bygone ideas or time for reinvention?" (Paper presented at the 29th Australasian Tax Teachers Association Conference, Masterton, Victoria University of Wellington, 2017).

⁵ N Faridy, B Freudenberg and T Sarker, "They are a tax fearing people: Deterrent Effect – Penalties, Audit and Corruption in a Developing Country" (2016) 3 *Curtin Law and Taxation Review* 1.

⁶ B.R. Jackson and V.C. Milliron, "Tax compliance research: findings, problems and prospects" (1986) 5 *Journal of Accounting Literature* 125.

⁷ M Richardson and A.J. Sawyer, "A Taxonomy of the Tax Compliance Literature: Further Findings, Problems and Prospects" (2001) 16 *Australian Tax Forum* 137.

only considered more countries in their analysis, they also identified emerging areas of tax compliance research, including the role of tax preparers and inducements.

While there have been reviews of tax compliance studies, since 1997 there has not been a published comprehensive analyses of tax compliance studies following the Jackson and Milliron model considering what have been the emerging trends, and the extent that our knowledge has improved with respect to understanding tax compliance. Given the growth of international trade, technological developments and sharing of information it is timely for this analysis that builds upon the model of Jackson and Milliron, and Richardson and Sawyer. This is the motivation for undertaking this project.

The aim of this article is to provide a comprehensive overview of the different tax compliance factors (concepts) that had been researched for the past twenty years (from January 1998 to December 2017) since the seminal papers by Jackson and Milliron⁸ & Richardson & Sawyer.⁹ This current study is a sequel to these two earlier studies, and given the 20 year time span and the amount of data, this article is the first of a series of articles. It focuses on the ‘variables’, with the other forthcoming articles considering ‘methodology’ and ‘theories’, respectively. In some respects, this is a descriptive exercise, drawing into one place the contributions of hundreds of published academic studies since the work of Richardson and Sawyer. It also draws upon other published work principally by academics in the interim that have sought to further our understanding of taxpayer compliance (and non-compliance). It is our intention to determine what new ideas in terms of variables have emerged in the intervening period.

We reviewed tax compliance studies and present content analysis of the selected articles to map the different factors/concepts that have been researched from 1998 to 2017. In this way we can determine whether the tax compliance factors (variables) identified prior to 1998 remain important, and whether new tax compliance factors have emerged. Data was drawn from 713 articles obtained from 15 research databases pertaining to taxation, accounting, social sciences, law, economics, business, political

⁸ Jackson and Milliron, above n 6.


⁹ Richardson and Sawyer, above n 7.

science and psychology journals. Given the large body of data analysed in this study, 713 articles, a computer-aided program for content analysis, namely Leximancer, was used. Leximancer assists in visualising the frequency and co-occurrence of concepts and themes. It is important to note that these databases are unlikely to include: books and book chapters; work published by major organisations (such as the OECD, World Bank and IMF); nor unpublished studies such as working papers.

This study offers a starting point, principally for academics, to be directed as to what has been researched and what needs to be researched further to extend our understanding of tax compliance in the new millennium. Also it provides those in tax policy development and analysis with a comprehensive foundation of our understanding of tax compliance from the perspective of academic research published in journals.

In considering the tax compliance research over the last twenty years it appears that while the 19 variables previously identified continue to be considered in tax compliance studies, new variables/concepts have emerged. In particular, there are 19 new variables emerging including tax evasion, tax enforcement, trust, power and sole traders. This analysis also demonstrates an apparent connectivity between some of the concepts, such as that of tax evasion with the concepts of perceived opportunity, tax avoidance and tax perception.

Additionally, this analysis provides some insights into the gaps in the tax compliance literature, as well as providing a pool of resources for new and experienced tax compliance researchers. This thorough analysis enriches our understanding of tax compliance. The findings suggest not only that our understanding of tax compliance (and non-compliance) has increased appreciably, they also build a strong foundation for future research as we continue to explore tax compliance in the new millennium.

Section Two of this article provides a broad summary of the prior studies conducted by Jackson & Milliron and latterly Richardson & Sawyer. The third section then provides an outline  of the research methodology undertaken and explains how Leximancer can be used for data analysis. Section Four sets out the results and our analysis. Through the analysis of the results, recommendations are proposed, with future research outlined in Section Five before our concluding comments.

Section Two: Literature Review

When a study is itself a literature review, then the ‘literature review’ could be viewed as being redundant. In our view it remains important to provide an overview of the prior literature reviews that have offered insights into what we know (and do not know) about taxpayer compliance.

Over time there has been a greater sophistication of our understanding of compliance, including cross-disciplinary research. For example, while income levels may be related to tax compliance levels, it may be that it has more to do with ‘types’ of income, as salary and wage earners have less opportunity to evade due to the withholding of taxes compared to other types of income (such as capital income). To consider how our understanding of income tax compliance has developed the two seminal papers by Jackson & Milliron,¹⁰ and then Richardson & Sawyer,¹¹ are discussed.

Jackson & Milliron: 1970s to 1985

In an attempt to aggregate our understanding of tax compliance, Jackson & Milliron¹² reviewed 43 income tax compliance studies which mainly focused on the USA. Through their review they synthesised 14 variables and four methodological approaches from the studies examined. Jackson and Milliron’s study focused on three main aspects of tax compliance research, namely:

1. The relationship between taxpayer compliance and 14 compliance variables frequently referred to in the literature (namely, age, gender, education, income level, income source, occupation, peer influence, ethics, fairness, complexity, revenue authority contact, sanctions, the probability of detection and tax rates);
2. Methodological issues associated with the four main research methods in the tax compliance field (namely surveys, experiments, analytic models and regression analysis);

¹⁰ Jackson and Milliron, above n 6.

¹¹ Richardson and Sawyer, above n 7.

¹² Jackson and Milliron, above n 6.

3. Four theoretical areas with potential relevance to future tax compliance research (namely prospect theory, deterrence theory, cognitive structures theory and agency theory).

By collating the findings of the 43 individual studies published prior to 1986, Jackson and Milliron sought to improve the understanding of taxpayer compliance as a whole, and identify those areas most in need of research attention in the future. Jackson and Milliron's definition of tax compliance as the reporting of all incomes and payment of all taxes by fulfilling the provisions of laws, regulations and court judgments has been the basis of many subsequent definitions of tax compliance. Jackson and Milliron's review is one of the most widely cited papers in the tax compliance field, and has been significant in guiding recent research in the area, including that of Richardson and Sawyer in 2001.

Jackson and Milliron's aggregation of studies demonstrated why there could be differences in research findings. For example, Jackson & Milliron, when discussing 'fairness', note that a 'more careful definition of the variable may be useful',¹³ since the cause of the differences in the research studies could be the different definitions and then how fairness is then operationalised in studies.

Jackson & Milliron observed that studies demonstrate that those who use professionals to help prepare tax returns can have greater audit adjustments, as they are more aggressive with their tax planning and use of 'grey' areas.¹⁴ Jackson & Milliron concluded their study with the observation that 'the state of tax compliance understanding can only be described as primitive', and that 'existing work should be examined critically for flaws'.¹⁵ Refer to Table 1 (later on in the article) for further information on the variables, research methodology and theoretical frameworks researched from the 1970s to 1985.

Richardson & Sawyer: 1986 to 1997

Richardson & Sawyer then examine the ten (10) years after the Jackson & Milliron study. Due to technological developments and increased activity in the area, they

¹³ Jackson and Milliron, above n 6, at 138.

¹⁴ Jackson and Milliron, above n 6, at 144.

¹⁵ Jackson and Milliron, above n 6, at 156.

analysed more articles (approximately 170 studies), and had a greater global mix of studies beyond just the USA.¹⁶ They suggested that the increased interest in tax compliance could in part relate to the political difficulty of raising tax rates (or to decrease public spending); and thereby if jurisdictions can improve tax compliance they can raise more tax revenue without awkward political decisions.¹⁷

To assist in comparison, Richardson & Sawyer highlighted the importance of actually defining what is meant by ‘compliance’, as well as the difference between ‘intentional’ and ‘non-intentional’ non-compliance. Richardson & Sawyer¹⁸ favoured the definition of tax compliance by Roth, Scholz and Witte:

“Compliance with reporting requirements means that the taxpayer files all required tax returns at the proper time and that the returns accurately report tax liability in accordance with the Internal Revenue Code, regulations, and court decisions applicable at the time the return is filed.”¹⁹

However, they observed that many of the studies analysed did not use this definition and some failed to have any definition at all. In this regard, Wu undertook a review of published research that examined ‘tax compliance’ to ascertain the degree to which there is an ‘agreed’ common understanding or definition of tax compliance.²⁰ Wu observes that there is a continuum with respect to the degree of severity of non-compliance, ranging from the unintentional to various degrees of intentional non-compliance. She concludes:

“The number of studies conducted around the area of behavioural tax compliance and non-compliance has led to a number of different versions of definitions for the relevant concepts; therefore, it is an extremely difficult task to attribute ‘universal’ definitions for each of the concepts at issue. Nevertheless, a reasonable conclusion can be drawn that ‘tax compliance’ *involves taxpayers satisfying the applicable rules and regulations by paying the correct amount of taxes due at the right time*. In addition, there is enough evidence to support the argument on the *importance to distinguish between taxpayers who observe the ‘letter’ of the law, the ‘spirit’ of the law, or both, to determine the ‘degree’ or*

¹⁶ Richardson and Sawyer, above n 7.

¹⁷ Richardson and Sawyer, above n 7, at 138.

¹⁸ Richardson and Sawyer, above n 7, at 142.

¹⁹ J.A. Roth, J.T. Scholz and A.D. Witte. (eds). *Taxpayer compliance: an agenda for research*. Vol. I. Philadelphia, PA (University of Pennsylvania Press, 1989), at 21.

²⁰ R Wu, “A Study on the Appropriateness for Adopting ‘Universal’ Definitions for Tax Compliance and Non-Compliance: A New Zealand Case Study Approach”, A thesis submitted in partial fulfilment of the requirements for the Master of Commerce Degree in Accounting and Information Systems at the University of Canterbury (2011).

‘extent’ of their compliance behaviour. Although it is again a subjective issue in ascertaining what is intended to mean by complying with the ‘spirit’ of the law, it may be reasonably justified that taxpayers who are ‘committed’ to meeting their tax obligations without enforcement activities by the tax authorities could fit into this particular category.”²¹

The matter of whether there should be a consistent definition for tax compliance remains unresolved. However, as a starting proposition, we agree that all studies should clearly set out the definition of tax compliance that they have adopted.

Richardson & Sawyer, in terms of their analysis of the 10 years since Jackson and Milliron’s study, observed that:

“... some variables have received significantly greater attention since 1985 (ethics, fairness, income level and the probability of detection), while others have not (occupation and revenue authority contact).”²²

In terms of our understanding how these variables may influence taxpayer compliance they concluded:

“Although the certainty associated with some variables has increased (revenue authority contact and sanctions) or been maintained (withheld income source, ethics and the probability of detection), the certainty associated with other variables has decreased (peer influence and tax rates).”²³

With the studies they analysed, Richardson & Sawyer, compared to Jackson & Milliron, found an additional five (5) variables (Compliance costs, Tax preparers, Framing, Positive inducements, and Tax amnesties); one (1) new research method (Process Tracing), and one (1) new research theory (Contingency Theory). Table 1 contains a summary of the Jackson & Milliron and the Richardson and Sawyer studies.

Richardson & Sawyer’s analysis highlights the multi-dimensional relationship between variables, such as gender and education effects on compliance. They also note that research investigating non-compliance by peers brings into question the direction of causality on compliance of various variables.

One of the conclusions by Richardson & Sawyer (2001, p 224) was that attention needs to be:

²¹ Wu, above n 20, at 147-148 (emphasis added).


²² Richardson and Sawyer, above n 7, at 203.

²³ Richardson and Sawyer, above n 7, at 203.

“... directed towards studies which develop and test their own hypotheses. ... as it is only through the replication of research findings (across populations, research methods, and time) that an increased level of certainty can be obtained about tax compliance behaviour.”²⁴

They also recommended that future non-compliance studies ‘examine separately the two main forms of taxpayer non-compliance — non-compliance through the overstatement of deductions, and non-compliance through the understatement of income’.²⁵

Considering the other side of the coin, Richardson & Sawyer observe that withholding and third party information reporting appears to improve compliance, although this needs to be balanced with possible adverse increases in tax compliance costs.²⁶ They also suggest that future studies need to delineate more for ethics, given that it appears the ethics of the taxpayer will affect the factors that motivate them. For instance, a less ethical taxpayer could be more influenced by the probability of detection, compared to taxpayers with high ethics, who could comply regardless.²⁷ However, the difficulty they highlighted with this was exactly how to identify and ‘measure’ ethics.²⁸

Refer to Table 1 for a summary of the tax compliance factors, research methodology and theoretical frameworks researched from 1986 to 1997. 

Since Richardson & Sawyer’s published study other researchers have provided extensive ‘partial’ reviews of the literature. Each has taken a different approach to that of Jackson and Milliron and Richardson and Sawyer. See, for example, work by Marjorie Kornhauser, conducted initially for the National Taxpayer Advocate (NTA) of the IRS (and subsequently published by the NTA) and separately by Kornhauser in the first decade of the 21st century.²⁹ James Alm and colleagues have also undertaken

²⁴ Richardson and Sawyer, above n 7, at 224.

²⁵ Richardson and Sawyer, above n 7, at 224.

²⁶ Richardson and Sawyer, above n 7, at 168.

²⁷ Richardson and Sawyer, above n 7, at 178.

²⁸ Richardson and Sawyer, above n 7, at 179.

²⁹ M Kornhauser, “Normative and Cognitive Aspects of Tax Compliance: Literature Review and Recommendations for the IRS Regarding Individual Taxpayers”, *Taxpayer Advocate Service, 2007 Annual Report to Congress* (2008, Volume Two), at 138-180. This was reprinted as: M Kornhauser, “A Tax Morale Approach to Compliance: Recommendations for the IRS” (2007) 8 *Florida Tax Review* 599.

a comprehensive analysis,³⁰ as has the International Monetary Fund (IMF).³¹ Other contributions include those of Borrego et al.,³² Devos;³³ and Khlif and Achek.³⁴ Also a number of other studies have included relevant sections of the tax compliance literature as part of their empirical analysis. A very recent review is that conducted by Ahmad Al-Zaqeba, Abdul Hamid, and Muhammad.³⁵

³⁰ See for example, J Alm, E Kirchler, S Muehlbacher, K Gangl, E Hofmann, C Kogler, and M Pollai, "Rethinking the Research Paradigms for Analyzing Tax Compliance Behaviour" (2012) 13(2) *CESifo Forum*, (Summer, 33-40. J Alm, E Kirchler and St Muehlbacher, "Combining Psychology and Economics in the Analysis of Compliance: From Enforcement to Cooperation" (2012) 42(2) *Economic Analysis and Policy* 133-151; and J Alm, "What Motivates Tax Compliance?", (2019) 33(3) *Journal of Economic Surveys*, 353-388

³¹ IMF, Current Challenges in Revenue Mobilization: Improving Tax Compliance (IMF, April 2015).

³² A.C. Borrego, C.M. da Mota Lopes and C.M. Santos Ferreira, "Tax Noncompliance in an International Perspective: a Literature Review" (2013) 14 *Portuguese Journal of Accounting and Management* 9.


³³ K Devos, *Factors Influencing Individual Taxpayer Compliance Behaviour* (Springer, Dordrecht, 2014).

³⁴ H Khlif and I Achek, "The determinants of tax evasion: a literature review" (2015) 57(5) *International Journal of Law and Management* 486.

³⁵ M Ali Ahmad Al-Zaqeba, S Abdul Hamid, and I Muhammad, Tax Compliance of Individual Taxpayers: A Systematic Literature Review, *Proceedings of the IIER International Conference*, Istanbul, Turkey, 22nd-23rd April 2018.

Table 1: Aggregated Tax Compliance Studies

Study	Period covered	Number and type of papers	Jurisdictions and study focus	Databases/Source	Variables identified		Methodologies identified	Theories identified
Jackson & Milliron (1986)	1970's to 1985	43 Including working papers and unpublished studies	USA (+ few studies outside USA) Compliance studies of income tax	Not specified	14 Variables described: 1. Age 2. Gender 3. Education 4. Income level 5. Income source 6. Occupation	7. Peer Influence 8. Ethics 9. Fairness 10. Complexity 11. Revenue Authority Contact 12. Sanctions, 13. Probability of Detection 14. Tax Rates	4 Methodological Approaches 1. Surveys 2. Experiments 3. Analytical models 4. Regression analysis	4 Theories identified: 1. Prospect Theory 2. Deterrence Theory 3. Cognitive Structures Theory 4. Agency Theory
Richardson & Sawyer (2001)	1986 to 1997 (10 years)	Approx 170 Only published studies	Multiple-jurisdictional Compliance studies of income tax (inc CGT if CGT include in 'assessable income')	4 databases identified: 1. ABI/Inform 2. EconLit 3. Social Sciences 4. Westlaw Also personal collection. Checked bibliography of gathered articles Physical checks of leading international journals	19 Variables described J & M's 14 Variables: 1. Age 2. Gender 3. Education 4. Income level 5. Income source 6. Occupation 7. Peer Influence 8. Ethics 9. Fairness 10. Complexity	11. Revenue Authority Contact 12. Sanctions, 13. Probability of Detection 14. Tax Rates And then 5 additional compliance variables: 15. Compliance costs 16. Tax preparers 17. Framing 18. Positive inducements 19. Tax amnesties	5 Methodological Approaches J & M's 4 Approaches: 1. Surveys 2. Experiments 3. Analytical models 4. Regression analysis And then 1 additional methodology: 5. Process Tracing	5 Theories identified J & M's 4 Theories: 1. Prospect Theory 2. Deterrence Theory 3. Cognitive Structures Theory 4. Agency Theory And then 1 additional theory: 5. Contingency Theory

Overall, the studies by both Jackson & Milliron  and Richardson and Sawyer provide an excellent summary and analysis of income tax compliance studies from the 1970s through to the end of 1997. However, since that time there has been continued interest in and need for research on tax compliance. Additionally, with the advent of technology, global trade and changing work and investment habits, it is important for us to understand what the studies since 1998 illuminate in terms of taxpayer compliance. The purpose of this study is to add to our understanding of tax compliance from published studies from 1998 through to 2017, with a particular focus on the *variables* identified. Given the magnitude of the data set, the other aspects, namely methodologies and theories, associated with tax compliance studies for the review period, are to be the subject of future articles.

Section Three: Methodology

This section provides the justification and the process involved in the Leximancer data analysis, and discusses the settings chosen for the purpose of the research. The prior studies of Jackson & Milliron, and Richardson & Sawyer, use a largely manual approach to their analysis. Given the size of the data set, Leximancer, was thought appropriate to assist in the initial analysis.

Leximancer

Leximancer is a text mining software that can be used to analyse the content of a collection of documents, as well as visually display the extracted analysis through conceptual maps.³⁶ These conceptual maps present the data from the main concepts and how these concepts relate to each other. To undertake this analysis, Leximancer uses empirically validated mathematical algorithms to make two determinations.³⁷ First, it determines the most frequently used concepts within a body of text (thematic analysis). Secondly, and more importantly, it determines whether there are relationships between these identified concepts (semantic analysis).³⁸ By using a

³⁶ See further: <https://info.leximancer.com/>.

³⁷ N Metropolis, A Rosenbluth, M Rosenbluth, A Teller and E Teller, "Equations of state calculations by fast computing machines" (1953) 21(6) *The Journal of Chemical Physics* 1087.

³⁸ M-L Verreynne, P Parker and M Wilson, "Employment systems in small firms: A multilevel analysis" (2011) 31(4) *International Small Business Journal* 405.

mathematical approach it produces ‘unbiased’ results by performing associational analysis of textual data in an independent manner.³⁹

Smith and Humphreys have demonstrated several advantages of Leximancer including that it can have face validity, stability, reproducibility, correlative validity and functional validity.⁴⁰ It achieves this by employing:

“proximity value to analyse text and learns to automatically identify and map theme and concepts in textual data. The software uses word frequency and co-occurrence data to identify families of terms that tend to be used together in the text. Thus, Leximancer identifies what concepts exist in a set texts, allowing concepts to be automatically coded in grounded fashion.”⁴¹

In this process:

“Leximancer bootstraps an expanded list of associated terms that signify a concept from the textual data. This boosts the recall of relevant contents, and more closely resembles the work conducted by a human coder”.⁴²

Important terms to understand in the use of Leximancer are: ‘concepts’, ‘themes’, ‘conceptual maps’, and within these ‘circles’ and ‘lines’; each of which is discussed below.

First, ‘concepts’ is the term used to describe words that occur very frequently. However, in considering these ‘concepts’, Leximancer goes beyond merely counting words, and instead uses coding blocks of sentences (generally either one or two sentence blocks). In this way, Leximancer allows us to go beyond a simple keyword search,⁴³ and can provide a co-occurrence factors which indicates which concepts are occurring with other concepts at the same time. A stronger co-occurrence indicates that these variables/factors (concepts) have a strong relationship, and they are usually investigated together. It also allows for compound words to be considered such as “tax evasion” and “tax morale”. For consistency with the prior studies, the terms ‘variables’

³⁹ P.W. Liesch, L. Hakanson, S.L. McGaughey, S. Middleton and J. Cretchley, “The evolution of the international business field: a scientometric investigation of articles published in its premier journal” (2011) 88(1) *Scientometrics* 17, at 24.

⁴⁰ A. Smith and M. Humphreys, “Evaluation of unsupervised semantic mapping of natural language with Leximancer concept mapping” (2006) 38 *Behavior Research Methods* 262.

⁴¹ J. Cretchley, D. Rooney and C. Gallois, “Mapping a 40-Year History with Leximancer: Themes and Concepts in the Journal of Cross-Cultural Psychology” (2010) 41(3) *Journal of Cross-Cultural Psychology* 318, at 319.

⁴² Liesch, Hakanson, McGaughey, Middleton and Cretchley, above n 39, at 24-25.

⁴³ Smith and Humphreys, above n 40, at 262.

and ‘factors’ will be used as a synonym for ‘concepts’ identified by Leximancer. Secondly, the term ‘theme’ refers to a group of related concepts which is ascertained by Leximancer through its coding blocks of sentences. A ‘theme’ will contain concepts that co-occur often within the same coding block,⁴⁴ with a theme represented by a circle on the concept map.

Through its analysis, Leximancer can then present ‘conceptual diagrams’ that present ‘concepts’, ‘themes’ and how they relate to each through the use of dots, circles and lines. The conceptual diagram allows for the interpretation of data in a pictorial form which would otherwise be too cluttered in a pure numerical form. The concept maps have ‘themes’ which are represented by a circle around several concepts. Each theme is named after the ‘most prominent concept in that group, which is also indicated by the largest dot in the theme cluster. A researcher can also rename themes when appropriate’.⁴⁵ The closeness between concepts indicates their co-occurrence in the data, and concepts that are more central in the circle indicates the extent of their connectivity to other concepts.⁴⁶

The size of the dot for each concept represents its prominence. The distance between the concepts within each circle also represents the proximity of concepts. The concepts are clustered according to weight and relationship. This means that concepts that appear together frequently in the data set or in similar situations are close together on the map. The grey lines show the most likely connections between concepts.


The advantage of Leximancer is that the Leximancer software can either run via default or can be directed to target key topics, as well as compare groups of different texts. As identified by Cretchley et al.:

“Using the concept map, rank-ordered concept lists and text query options, the researcher can read instances of specific concepts to understand their relationships. Thus, the map visually represents the strength of association between concepts and provides a conceptual overview of the semantic structure of the data.”⁴⁷

⁴⁴ Cretchley, Rooney and Gallois, above n 41, at 319.

⁴⁵ Cretchley, Rooney and Gallois, above n 41, at 319.

⁴⁶ Verreynne, Parker and Wilson, above n 38.

⁴⁷ Cretchley, Rooney and Gallois, above n 38, at 319. 

Who has used Leximancer?

Leximancer has been applied for various research purposes, including: literature reviews, coding open-ended surveys, coding interview transcripts, analysing media articles, analysing website content (e.g. a university site profile), and other text documents, such as submissions from stakeholders. Table 2 summarises the common applications of Leximancer in accounting and management studies.

Leximancer has been used for data analysis in various social and business- related academic research, including psychology,⁴⁸ accounting,⁴⁹ legal duties,⁵⁰ management disciplines,⁵¹ and more recently taxation.⁵²

⁴⁸ Cretchley, Rooney and Gallois, above n 41.

⁴⁹ M Axelsen, P Green and G Ridley, "Explaining the information systems auditor role in the public sector financial audit" (2017) 24 *International Journal of Accounting Information Systems* 15; K Crofts and J Bisman, "Interrogating accountability: An illustration of the use of Leximancer software for qualitative data analysis" (2010) 7(2) *Qualitative Research in Accounting and Management* 180; P.D. Palmer, "Exploring attitudes to financial reporting in the Australian not-for-profit sector" (2013) 53(1) *Accounting & Finance* 217.

⁵⁰ T Voogt and M-L Verreynne, "Director Appointments: Expressing Board Care and Diligence" 41(4) *UNSW Law Journal* 1335.

⁵¹ Verreynne, Parker and Wilson, above n 35; M Ho, D Nguyen, K Lo, C McLean and S Teo, "Research Note: The state of human resource (HR) competency research: Charting the research development of HR competencies and examining the signals from industry in New Zealand" (2015) 40(3) *New Zealand Journal of Employment Relations* 67; B McKenna, M-L Verreynne and N Waddell, "Locating gendered work practices: a typology" (2016) 37(6) *International Journal of Manpower* 1085; W Cukier, S Gagnon, E Roach, M Elmi, M Yap and S Rodrigues, "Trade-offs and disappearing acts: shifting societal discourses of diversity in Canada over three decades" (2017) 28(7) *The International Journal of Human Resource Management* 1031; M Maluwetig and M-L Verreynne, Entrepreneurship education: How socially focused should we be? (Paper presented at the Swinburne University of Technology. Symposium conducted at the meeting of the 8th AGSE International Entrepreneurship Research Exchange, Melbourne, Australia, 1-4 February 2011).

⁵² B McCredie and K Sadiq, An Examination of the Legitimacy of the Aggregate View of the Corporation (Paper presented at the 30th Australasian Tax Teachers Association Conference, Monash University, Melbourne, 17-19 January 2018).

Table 2: Applications of Leximancer in Academic Research Settings

Application	Research Purpose	Type of Text	Output Options	Authors
Literature Reviews	To investigate the contextual concepts of accountability in accounting journals.	Analysed 114 journal articles related to accountability published in highly ranked accounting journals from 2000 to 2017.	Visual via concept maps and report via statistical output.	(Crofts & Bisman, 2010) – accounting discipline
	To track the development of themes and concepts within the Journal of Cross-Cultural Psychology over 40 years.	Analysed 1416 articles mapped by decade.	Concept maps and representative quotes from text extracts.	(Cretchley et al., 2010) – psychology discipline
	To chart the movement of HR competency research and the signals given by organisations in the recruitment of HR professionals in New Zealand	Analysed 51 HR competency research articles from the Business Source Premier database from 1990-2012.	Concept maps, Insight Dashboard and statistical outputs.	(Ho et al., 2015) – management discipline
Coding Open-ended Surveys	To explore the perspectives of employees and managers from 50 small Australian firms to build a more complete and nuanced view of the interaction of human resources and capabilities with firm performance.	Analysed the qualitative open text data from six questions in the CEO survey and 30 questions in the employee survey.	Concept maps and representative quotes from text.	(Verreynne et al., 2011) – management discipline
Coding Interview Transcripts	To explore how work and life are discursively constructed by men and women in their daily practice of living	Analysed data from focus group and semi-structured interviews of 18 female and 19 male Australian managers in the financial and government sector	Concept maps and statistical output.	(McKenna et al., 2016) – management discipline
	To identify relevant “common statements” representing the prominent and shared perceptions of the IS auditor role amongst these auditors.	Analysed 32 interview transcripts with IS auditors on their roles in the public sector financial audit	Statistical output and representative quotes from text.	(Axelsen et al., 2017) – accounting discipline
Coding Media Articles	To explore the discursive constructions of equity and diversity in Canada at societal/macro level and provides original findings on how these discourses have been produced in the public sphere	125 articles published from 1986 to 2014 in The Globe and Mail (Canadian newspaper) and 26 report from the Canadian Employment Equity Act Annual Reports from 1988 to 2013.	Concept maps, statistical output and representative quotes from text.	(Cukier et al., 2017) – management discipline
Coding Website Content	To investigate the prevalence of social dimensions in business education in high-performing universities and to examine the relationship between their ranking and	Text data extracted from the university’s website of the top-ranking universities to reveal preferences, priorities, and	Concept maps	(Maluwetig & Verreynne, 2011) – management

Application	Research Purpose	Type of Text	Output Options	Authors
	social dimensions	recurring themes among high performers.		
Stakeholder Relationship Management	To explore attitudes to financial reporting in the not-for-profit (NFP) sector, by utilising submissions to the 2008 Australian Senate Economics Standing Committee.	Examined submissions from NFP entities and interested parties to the Senate	Concept Map and representative quotes from text.	(Palmer, 2013) – accounting
	To illustrate the complementary use of two computational semantic analysis techniques – latent semantic analysis and data mining - to examine a large textual data set.	Analysed 8544 abstracts from selected journals in Information Systems, Management, and Accounting over a period of 25 years.		(Indulska, Hovorka, & Recker, 2012) - Management and Accounting
Coding Roles and Responsibilities	To determine the role and responsibilities of company directors from the codes and/or legislation from five jurisdictions: UK, Singapore, Australia, G20/OECD and South Africa.	Codes and legislative instruments governing company directors to determine key concepts and themes in their role and responsibility.	Concept Map and representative quotes from text.	(Voogt & Verreynne, 2018) – Business law.

Of course, with any tool there are advantages and disadvantages. A clear advantage of Leximancer is its ability to effectively analyse a large quantity of text data, and for the purposes of this study this involved over 700 articles. However, for data (articles) to be analysed they need to be in ‘editable’ text format; this means that either they are in a Word document or editable PDF document. While it is possible to convert a read only PDF into Word, this may lead to some formatting errors. For example, the documents created may contain sentences truncated by paragraph marks, words containing unreadable symbols or incorrect spacing between words, which may result in inaccurate coding of concepts. Although it would be ideal to correct all these text processing errors before inputting the files into Leximancer, this process requires manual editing which is extremely time-consuming and prone to human error.⁵³

While Leximancer allows for the rapid identification of concepts,⁵⁴ some of the data (words) can be irrelevant, such as headers and footers, headings, references, that can appear in the formatting of an article. This requires manual removal of irrelevant text content (e.g., headings and references) prior to data analysis. While this is time consuming it provides for richer data for content analysis.

A clear advantage of Leximancer is objective data analysis as it avoids problems of researcher bias coder subjectivity.⁵⁵ However, this analysis only provides a ‘birds-eye’ conceptual overview of the text data only. Articles that focus on specific concept(s)

⁵³ In some converted files, sentences that span several lines may be truncated by paragraph marks, resulting in inaccurate coding of concepts and their co-occurrence with other concepts within the same sentence. For example see one by M Burton, “Citizens as Partners? Foundations for an Effective Tax System in the New Democratic Era” (2006) 24(2) *Tax Law and Political Institutions* 169, at 171:

“What much of this literature misses, however, is that the same scepticism which undermines absolute norms of substantive fairness also undermines any claim that a particular political process, such as deliberative democracy, is the ‘right’ process (Bohman, 1998; Mouffe, 1995).”

In this example Leximancer treated this quote as two sentences as there is a paragraph mark before the word ‘scepticism’, which led to a commencement of a new sentence.

There is a second example by D DeÁK, “Legal Considerations of Tax Evasion and Tax Avoidance” (2004) 26(1) *Society and Economy* 41, at 43:

“In reviewing tax avoidance schemes, it is essential to identify the inherent logic and the authentic content of a legal system. The validity of law is a category equivalent to the micro category of the coherence of law.”

This example shows that although there is a missing space between the first and second sentences Leximancer has still recognised them as two sentences because of the full-stop after the word ‘system’. However, the phrase ‘of law’ is treated as one single word due to a missing space.

⁵⁴ Smith and Humphreys, above n 40; and P Sotiriadou, J Brouwers and T.A. Le, “Choosing a qualitative data analysis tool: a comparison of NVivo and Leximancer” (2014) 17(2) *Annals of Leisure Research* 218.

⁵⁵ Smith and Humphreys, above n 40.

will have higher frequency counts for those concepts due to the narrow focus of the article (particularly in the results and discussion sections of the same article where the key concept(s) are mentioned repeatedly). Conversely, a literature review article will have more widespread frequency counts for a wider array of concepts.

Part of Leximancer's reliability is that little researcher intervention is required in content analysis process,⁵⁶ and text segments are always coded in the same way given the same parameter settings.⁵⁷ However, there can be some issues for Leximancer concerning jargon and similar terms. For the purposes of our study, we disabled the self-learning thesaurus and prescribed multiple meanings for certain concepts. For instance, the concept 'tax agents' is also referred to as 'tax advisors', 'tax consultants', 'tax lawyers' and 'tax practitioners' in the literature. However, this can lead to concepts in a compound term being implicitly stated anywhere in the text segment. For instance, the compound concept 'tax rate' was also coded where the term 'rate' was used to denote 'evasion rate', 'audit rate', 'penalty rate', and 'compliance rate'. We did not isolate coding of other terms for this compound concept, as these concepts were also important to the study.

Also, the concept diagrams themselves are stochastic, meaning that the concepts on the map may settle in different positions each time a new map is generated from the software. This is because the concepts were initially distributed randomly in multidimensional space.⁵⁸ However, any instability in the map does not affect the frequency statistics or the centrality rankings of concepts.⁵⁹

Overall, it was determined that it was appropriate to use Leximancer as a semantic mapping methodology to organise the large amount of literature on tax compliance. Leximancer could also be used to explore the conceptual complexity of tax compliance factors (variables) and the interrelationship between the factors. How this process occurs is described in the next subsection.

⁵⁶ Sotiriadou, Brouwers and Le, above n 54.

⁵⁷ Smith and Humphreys, above n 40.

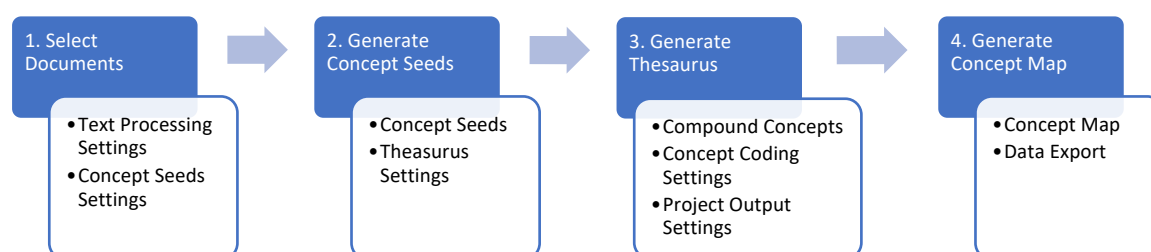
⁵⁸ Leximancer Pty Ltd. (2016). *Leximancer User Guide Release 4.5* Retrieved from info.leximancer.com), at 89.

⁵⁹ Smith and Humphreys, above n 40.

Data processing and data analysis

Leximancer (version 4.5) was used for this study, and Figure 1 illustrates the four key processing stages in the Main Leximancer User Interface: 1. Select Documents, 2. Generate Concept Seeds, 3. Generate Thesaurus and 4. Generate Concept Map.

Figure 1: Leximancer Project Processing Stages



Source: adapted from the Main Leximancer User Interface (<http://localhost:8085/leximancer/>)

In brief, the Select Documents stage enables the user to select the source of documents for analysis and specify configuring parameters for text processing and concept seeds.

In terms of the Select Documents, studies on tax compliance factors were researched and developed from 1 January 1998 to 31 December 2017. Due to the multidisciplinary nature of tax compliance, articles were gathered from taxation, accounting, social sciences, law, economics, business, political sciences and psychology journals. Fifteen databases were used to access these journals, and they are listed in Table 3.

Table 3: Databases used



ABI-Inform Complete (Proquest) Accounting and Tax (Proquest) Business Source Complete ScienceDirect AGIS Plus Text Australian Taxation Law Library Google Scholar	Heinonline JSTOR Kluwer Law Lexis Advance LexiNexisNZ Thomson Reuters WestLaw WestLaw NZ and Australia BestCase (via WestlawNext Canada).
---	--

The breadth of the current study appears to be more comprehensive than the prior two studies given the number of databases used to draw research from (refer to Table 1 earlier).

Within these databases, keywords used for the search were “tax* compliance” (truncation is used for ‘tax’ to include taxation and taxpayer) on peer reviewed, full-text, academic journal articles from 1998 to 2017 inclusively. It is acknowledged this approach introduces a major limitation. A large part of the compliance literature has been produced by institutions such as World Bank, Organisation for Economic Cooperation and Development (OECD), IMF, Internal Revenue Service (IRS) and the like, and it is arguable that this literature has had more influence than much of the academic literature. The current study extended beyond ‘income tax’ compliance studies that was the single focus of Jackson & Milliron⁶⁰ and Richardson & Sawyer.⁶¹ Specifically, it includes VAT and GST compliance studies, although studies about fuel tax, inheritance tax and land tax were excluded, since these are typically less significant taxes. Also, limited search query (e.g. abstract only) and additional keywords (e.g. AND factor OR determinant OR variable; AND gender OR culture OR religion OR religiosity) were used for Kluwer Law, Thomson Reuters WestLaw and WestLaw NZ databases to further refine the search results.

From this initial process a total of 913 articles were identified through the 15 databases. Some 132 articles were excluded as they were not accessible and/or were not relevant.⁶² A further 128 articles were also excluded as they related to tax policies and tax reform articles which were not relevant for the purpose of this research.⁶³ This resulted in a total of 713 papers being eligible for further text analysis. The 713 eligible papers were then divided into two main groups: conceptual and empirical. Conceptual papers are discussion papers without empirical data; whereas empirical

⁶⁰ Jackson and Milliron, above n 6.

⁶¹ Richardson and Sawyer, above n 7.

⁶² These articles came up in the initial search results but in the second stage of screening and sorting they did not fit into the research type criteria or they only mentioned tax compliance several times (less than 10 times and mainly within the reference/note section) in the paper, meaning that the focus of the paper is too broad and not suitable for the purpose of our research. A small number of papers were accessible but were only available in scanned and uneditable PDF format and could not be used for Leximancer analysis.

⁶³ Tax policies and tax reform articles were excluded because they did not focus on tax compliance variables but were more focused on legislative history, requirements and suggestions for reforms.

papers are supported by sample data. Conceptual papers were tagged under the Literature Review folder and the empirical papers were tagged under the Quantitative, Qualitative and Mixed Method folders. The names of the folders reflect the different methodological approaches used. Of the 713 articles, 174 were conceptual papers, and 539 were empirical papers. From the 713 articles, 260 papers were from 1998 to 2007 and 453 from 2008 to 2017. The significant increase in published tax compliance articles in the second decade demonstrates a greater level of activity in the topic, which could be due to the implications of tax compliance on tax revenues, policies and practice. Table 4 shows the breakdown between the conceptual and empirical groups and the research methods used of tax compliance articles selected for this study.

Table 4: Type of Research

Type of Research	1998-2017 (n=713)
Conceptual papers (tagged under Literature Review folder)	174
Empirical papers (tagged under Quantitative, Qualitative and Mixed Methods folders)	539
Quantitative Folder comprised of:	
- Survey	200
- Experiment	115
- Analytical Modelling	95
- Regression Analysis	75
Total Quantitative	485
Qualitative folder comprised of:	
- Interviews	20
- Case Study	8
- Focus Groups	1
- Observation	2
- Content Analysis	3
- Diary or Log Book Study	0
Total Qualitative	34
Mixed Methods	20

If nothing else, there is apparently growth in the research being conducted in this area, as 713 articles were obtained from 15 research databases compared to the 43 studies considered by Jackson & Milliron,⁶⁴ and approximately 170 studies in Richardson &

⁶⁴ Jackson and Milliron, above n 6.

Sawyer.⁶⁵ For this study this represents an average of 35 papers per year, compared to 14 papers per year for the ten years considered by Richardson & Sawyer.

In the text processing and concept seed setting we have adopted all default settings with the exception of setting the context blocks for coding as single sentences instead of using the default context blocks of two sentences. Single sentence context blocking was appropriate to ensure the context block was a cohesive statement and enables compound concepts (see stage 3 Generate Thesaurus) to be coded within a single sentence block. We tagged the document folders in the text processing setting to generate concept maps that identify and compare concept frequency and relationships between each research type (i.e., literature review, quantitative research, qualitative research and mixed methods research).

Although Leximancer allows the user to generate an automatic map by completing the Select Document stage followed immediately by the Generate Concept stage, we have also completed the Generate Concept Seeds and Generate Thesaurus stages with specified settings. In the Generate Concept Seeds stage, we removed all automated concept seeds and created a list of user-defined concept seeds instead. This is because our research focuses on the tax compliance literature which utilises a number of specific terminologies (which followed the terms used in Jackson & Milliron and Richardson & Sawyer - see Table 1 earlier). In addition, we disabled the self-learning thesaurus function and manually merged concepts with similar meanings (e.g., religion and religiosity).

For the Generate Thesaurus Stage we created compound concepts which were concepts comprising of more than one term (e.g., ‘tax’ AND ‘evasion’, ‘trust’ AND NOT ‘financial’ AND NOT ‘foreign’) appearing within a context block. We then selected the concepts to be included in the output maps. The settings chosen for each of the above mentioned processing stages are listed in Table 5 .

⁶⁵ Richardson and Sawyer, above n 7.

Table 5: Leximancer processing and analysis settings

Leximancer Settings	
Stage One: Select Document	
1.	Open the Select Document Settings, select folders containing source documents
2.	Open Text Processing Options, go to the General column - sentences per block, click on the drop down menu and select 1 to reduce corpus to 1-sentence
3.	Go to the Tab column, tick the box next to Folder, click OK, close the Text Processing Options screen
Stage Two: Generate Concept Seeds	
1.	Run the Generate Concept Seeds Process
2.	Open Concept Seed Editor, click on the Auto Concepts Tab, select all concepts and click Remove, click OK
3.	Click on the Auto Tags Tab, merge and/or edit folder names, click OR
4.	Click on the User Defined Concepts Tab, add concepts and merge concepts with similar meanings, click OK, close the Concept Seed Editor screen
5.	Open the Thesaurus Setting (Concept Learning), go to the General column – Learn Concept Thesaurus using Source Documents, untick the box to disable self-learning thesaurus, click OK, close the Thesaurus Setting screen
Stage Three: General Thesaurus	
1.	Run the General Thesaurus
2.	Open Compound Concepts, select name-like and word-like concepts lists in the first two columns, add to the right-hand column and create compound concepts using AND, OR and/or NOT functions, edit the names of compound concepts, click OK
3.	Open Concept Coding Settings, go to the Mapping Concepts Tab, select all concepts in the Selected Mapping Concepts column and click on the right-hand arrow button to remove them from the list, then select folder tags from the Available Names column and mapping concepts from the Available Concepts column and click on the left-hand arrow button to add them to the Selected Mapping Column, click OK
Stage Four: Generate Concept Map	
1.	Run Generate Concept Map
2.	Open Concept Map
3.	Adjust concept visibility, theme size and rotation
4.	Edit font size, theme and concept labels
5.	Save concept map
6.	Export concept map, theme summary, ranked concept lists and query results

The initial analysis yielded 107 terms, with *tax* occurring most frequently and *rights* the least. We enhanced the semantic usefulness of Leximancer findings by further refining the analysis process with the following steps:

1. We deleted concepts of little relevance to the paper (e.g. data, study) and name-like concepts (e.g. FATCA and IRD);
2. We merged concepts which are synonymous with others (e.g. cost and costs, businesses and companies);
3. We added concepts that were of interest to our study but not identified as frequent concepts by Leximancer (e.g. culture, religion);⁶⁶ and
4. We created compound concepts, which are comprised of two or more concepts that appear in the same sentence text (e. g. sole traders, tax morale).

Table 6 provides the thesaurus for the 38 tax compliance concepts. The thesaurus clarifies the tax compliance concepts currently reported in this study. It does so by reducing duplication of similar words such as ethics or ethical. The thesaurus also provides the derivation of compound words used such as income level and perceived opportunity. This is to provide transparency in the data processing for the 713 papers.

Table 6: Thesaurus for the 38 Tax Compliance Concepts

Tax Compliance Concepts	Thesaurus
Age	taxpayer AND age
authorities contact	authorities AND contact
business taxpayer	business AND taxpayer
Culture	cultural OR culture OR ethnic OR ethnicity OR race
Education	taxpayer AND education
Ethics	ethics OR ethical
Framing	frames OR framing
Gender	taxpayer AND gender
income level	income AND level
income source	income AND source
individual taxpayer	individual AND taxpayers
Occupation	taxpayer AND (occupation OR status)
peer influence	peer AND influence
perceived opportunity	perceived AND opportunity

⁶⁶ This means there appearance is because Leximancer was instructed to look for these concepts.

Power	Power
Political	Political
probability of detection	(probability AND detection) OR (tax AND audit)
Religion	religion OR religiosity
Reciprocity	Reciprocity
Sanctions	sanction OR sanctions
social norms	social AND norms
sole traders	self-employed OR SME(s) OR (sole AND (enterprises OR traders OR proprietors)) OR ((micro OR small) AND businesses)
tax agents	(tax OR taxation) AND (agent(s) OR accountant(s) OR advisor(s) OR consultant(s) OR lawyer(s) OR practitioner(s) OR preparer(s))
tax amnesty	(tax OR taxation) AND (amnesty OR amnesties)
tax avoidance	(tax OR taxation) AND avoidance
tax complexity	(tax OR taxation) AND complexity
tax compliance cost	(tax OR taxation) AND compliance AND cost
tax evasion	(tax OR taxation) AND evasion
tax enforcement	(tax OR taxation) AND enforcement
tax fairness	(tax OR taxation) AND fairness
tax inducement	((tax OR taxation OR positive) AND inducements) OR ((tax OR taxation) AND (penalty OR penalties))
tax morale	(tax OR taxation) AND morale
tax perception	(tax OR taxation) AND perception(s)
tax rate	(tax OR taxation) AND rate(s)
tax system	(tax OR taxation) AND system
Trust	trust AND (NOT financial) AND (NOT foreign)
wealthy taxpayer	(wealthy OR (high AND wealth)) AND taxpayer(s)
Withholding tax	(withholding tax*)

To facilitate an analysis of different time series in terms of how tax compliance research may have altered between two decades (first decade: 1998 to 2007; second decade: 2008 to 2017) we created subsets of data in 10-year periods.

The next section provides a summary of the results of the Leximancer analysis.

Section Four: Results

This article contributes to research on tax compliance in three ways. First, it collects and summarises multiple tax compliance studies over the last two decades from 1998 to 2017 inclusive. Second, it provides a snapshot overview of the different concepts and themes that are pertinent and researched since the Richardson & Sawyer article. Third, it compares the different themes and concepts that were researched between

those in the prior studies of Jackson & Millron and Richardson & Sawyer. There are three segments in this section which are:

1. Research concepts and themes identified from 1998 to 2017;
2. The top five themes researched over the last 20 years; and
3. Concepts researched in the period 1998 to 2017, compared to Jackson & Milliron and Richardson & Sawyer.

Research concepts and themes identified from 1998 to 2017

Through the analysis of the 713 articles, 38 variables (concepts) were identified, which means that an additional 19 variables have emerged since the prior studies by Jackson & Milliron and Richardson & Sawyer. Table 7 provides an overview of the results of this study, including the variables/concepts identified.

Table 7: Variables Identified 1998 to 2017

Study	Period covered	Number and type of papers	Jurisdictions and study focus	Databases/Source	Variables identified	
Sue Yong, Karen Lo, Brett Freudenberg, and Adrian Sawyer	Jan 1998 to Dec 2017 (20 Years)	713 Only published refereed articles (260 for 1998 to 2007) (453 for 2008 to 2017)	Multiple-jurisdictional Compliance studies which looked at income tax, CGT, VAT (GST). (Excluded compliance studies about fuel tax, inheritance tax and land tax)	15 databases utilised: 1. ABI-Inform Complete (Proquest) 2. Accounting and Tax (Proquest) 3. Business Source Complete 4. ScienceDirect 5. AGIS Plus Text 6. Australian Taxation Law Library 7. Google Scholar 8. Heinonline 9. JSTOR 10. Kluwer Law 11. Lexis Advance 12. LexiNexisNZ 13. Thomson Reuters WestLaw 14. WestLaw NZ and Australia 15. BestCase (via WestlawNext Canada).	38 variables described: J & M's 14 Variables: 1. Age 2. Gender 3. Education 4. Income level 5. Income source 6. Occupation, 7. Peer Influence, 8. Ethics 9. Fairness 10. Complexity 11. Revenue Authority Contact 12. Sanctions, 13. Probability of Detection 14. Tax Rates R & S's 5 additional compliance variables: 15. Compliance costs, 16. Tax preparers 17. Framing 18. Positive inducements 19. Tax amnesties Additional 19 variables (concepts) identified: 20. Tax evasion (more a theme than a factor) 21. Tax enforcement 22. Trust 23. Power 24. Sole traders 25. Culture 26. Tax System 27. Political	28. Individual taxpayer 29. Business Taxpayer 30. Tax morale 31. Social Norms 32. Religion 33. Withholding taxes 34. Reciprocity 35. Tax perception 36. Perceived opportunity 37. Tax avoidance 38. Wealthy taxpayer



These 38 concepts are associated with the key five themes of Tax Evasion, Tax System, Tax Agents, Tax Morale and Ethics. The way that the concepts are clustered around themes is detailed in Table 8 (with new additional concepts denoted by “*”).

Table 8: Tax Compliance Concepts contained in each Theme 

Themes	Tax Compliance Concepts contained in each Cluster	Number of Concepts
Tax evasion	1. Tax evasion* 2. Tax rate 3. Probability of detection 4. Trust* 5. Tax inducement 6. Tax fairness 7. Power* 8. Tax perception* 9. Tax enforcement* 10. Income level 11. Reciprocity* 12. Peer influence	12
Tax system	13. Tax system* 14. Tax avoidance* 15. Political* 16. Sanction 17. Tax complexity 18. Individual taxpayer* 19. Business taxpayer* 20. Income source 21. Withholding taxes* 22. Authorities contact 23. Wealthy taxpayer*	10
Tax agents	24. Tax agents 25. Tax compliance cost 26. Sole traders* 27. Culture* 28. Education	5
Tax morale	29. Tax morale* 30. Tax amnesty 31. Framing 32. Perceived opportunity*	4
Tax morale	33. Ethics 34. Social norms* 35. Religion* 36. Occupation 37. Age 38. Gender	6

Note:* denotes tax compliance variables identified in our study in addition to those examined by Jackson and Milliron⁶⁷ and Richardson and Sawyer.⁶⁸ See Table 1 and Table 7 for more detail.

⁶⁷ Jackson and Milliron, above n 6: examined fourteen compliance variables: age, gender, education, income level, income source, occupation, peer influence, ethics, fairness, complexity, revenue authority contact, sanctions, the probability of detection and tax rates.

⁶⁸ Richardson and Sawyer, above n 7: added five other variables: compliance costs, tax agents, tax amnesties, framing of tax decisions and positive inducements.

The additional tax compliance concepts (variables) researched since Richardson and Sawyer article are: Tax evasion; Tax enforcement; Trust; Power; Sole traders; Culture; Tax system; Political; Individual taxpayer; Business taxpayer; Tax morale; Social norms; Religion; Withholding taxes; Reciprocity; Tax perception; Perceived opportunity; Tax avoidance and Wealthy taxpayer. These issues were given more attention by researchers over the last twenty years. It could be that these terms are inter-related to the prior variables identified but are now more nuanced, or they could be new variables that have been explored. For example the new concept of ‘tax enforcement’ is likely to be related to the prior variable of ‘sanctions’.



This analysis is then presented as a diagram to demonstrate the relationships between different concepts and the type of research conducted, whether it be a literature review, quantitative, qualitative or mixed methods study. See further Figure 2.



Leximancer Concept Map – Tax Compliance Conceptual Papers and Empirical Research for 1998-2017 (visible concepts 100%, theme size 39%)

Additional information regarding the concepts is provided in Table 9, which ranks the concepts by count and relevance scores. This shows the number of times each concept has been mentioned within the 713 articles. As shown in Table 9, *tax evasion* has the highest count followed by *tax rate*, *tax system*, and *probability of detection*. On the other hand, the lowest counts were from *wealthy taxpayer* and *perceived opportunity* concepts with only 42 and 30 mentions respectively within the texts. While this does not indicate importance nor which factors (variables) have the greatest influence on tax compliance it does indicate the number of times concepts are mentioned. The prominence of ‘tax evasion’ could indicate that to understand ‘compliance’ we need to better understand why people evade their tax.⁶⁹ The relevance count would tend to suggest that the following concepts are more relevant to tax compliance in the new millennium as they have over a 30% count relative to the top concept: tax evasion; tax rate; tax system; probability of detection; tax morale; trust; tax agents and tax compliance cost. Whereas the following concepts appear weaker with a 5% or less count relative to the top concept: income source; framing; reciprocity; withholding taxes; education; occupation; age; gender; authorities contact; peer influence; wealthy taxpayer and perceived opportunity. However, it could be that these factors are less researched than the others (which may be due to methodological issues, or difficulties in measuring).

⁶⁹ This was the focus of earlier research: J Alm, GH McClelland and WD Schulze, “Why do people pay taxes?” (1992) 48 *Journal of Public Economics* 21-38.

Table 9: Ranked List of Tax Compliance Concepts by Count and Relevance Scores

Concepts	Count	Percentage of Count Relative to the Top Concept
tax evasion	8414	100%
tax rate	5476	65%
tax system	4363	52%
probability of detection	4292	51%
tax morale	3560	42%
trust	3465	41%
tax agents	2999	36%
tax compliance cost	2806	33%
sole traders	2439	29%
tax inducement	2197	26%
power	2116	25%
ethics	2072	25%
culture	1868	22%
tax avoidance	1814	22%
tax enforcement	1610	19%
political	1556	18%
tax fairness	1540	18%
sanction	1448	17%
tax perception	1414	17%
social norms	1269	15%
tax complexity	1042	12%
individual taxpayer	1032	12%
income level	1015	12%
tax amnesty	769	9%
business taxpayer	763	9%
religion	516	6%
income source	444	5%
framing	432	5%
reciprocity	328	4%
withholding taxes	301	4%
education	217	3%
occupation	138	2%
age	138	2%
gender	72	1%
authorities contact	65	1%
peer influence	56	1%
wealthy taxpayer	42	0%
perceived opportunity	30	0%

Note:

Count = the number of times the concept appears in the entire data set (per one-sentence block)

Percentage of Count Relative to the Top Concept (or relevance percentage) = the count value of each concept divided by the single highest count value. The relevance percentage for tax rate is calculated by taking the count value for tax rate (5476) then divide it by tax evasion's count value (8414) then times 100% = 65%. Thus, the most frequent concept will always have 100% relevance, regardless of whether it occurs in all context blocks.

Top Five Concepts

Table 10 ranks the tax compliance themes by count and connectivity scores. This information is important in determining the strength of the relationships (if any) between the concepts. It shows the connection of each concept to one another rather than the count relative to the most important concept which is tax evasion.

Table 10: Ranked List of Tax Compliance Themes by Count and Connectivity Scores



Theme	Concept	Count	Total Count	Connectivity	Total Connectivity
1. Tax evasion	tax evasion	8414	31923	4097	20266
	tax rate	5476		3266	
	probability of detection	4292		2807	
	trust	3465		1646	
	tax inducement	2197		1982	
	tax fairness	1540		1745	
	power	2116		1107	
	tax perception	1414		1677	
	tax enforcement	1610		995	
	income level	1015		709	
	reciprocity	328		108	
	peer influence	56		127	
2. Tax system	tax system	4363	12870	2502	6985
	tax avoidance	1814		934	
	political	1556		531	
	sanction	1448		615	
	tax complexity	1042		936	
	individual taxpayer	1032		422	
	business taxpayer	763		375	
	income source	444		360	
	withholding taxes	301		157	
	authorities contact	65		134	
	wealthy taxpayer	42		19	
3. Tax agents	tax agents	2999	10329	1248	3921
	tax compliance cost	2806		1030	
	sole traders	2439		706	
	culture	1868		703	
	education	217		234	
4. Tax morale	tax morale	3560	4791	1340	1710
	tax amnesty	769		252	
	framing	432		84	
	perceived opportunity	30		34	
5. Ethics	ethics	2072	4205	913	2266
	social norms	1269		593	
	religion	516		271	
	occupation	138		165	
	age	138		164	
	gender	72		160	
Total			64118		35148

Each concept has a count value and connectivity score.

Count value measures the number of times a concept appears within the data set. For instance, Table 10 shows that the concept *tax evasion* has a count score of 8414, meaning that this concept has been coded 8414 times in the text. Therefore, the total concept counts for the theme *Tax evasion* = tax evasion 8414 + tax rate 5476 + probability of detection 4292 + trust 3465 + tax inducement 2197 + tax fairness 1540 + power 2116 + tax perception 1414 + tax enforcement 1610 + income level 1015 + reciprocity 328 + peer influence 56 = 31923.

Connectivity score measures the number of times a concept appears concurrently with another concept within the same sentence in the data set. Note the ‘*’ highlights the new variables/concepts compared to the prior study of Richardson and Sawyer.

Figure 3 illustrates that when the concept *tax evasion* is selected in the concept map, a number of red radiant lines appear connecting this concept with various other concepts. The connectivity scores for this concept are displayed on the count column on the right. According to the connectivity scores, *tax evasion* is coded 13 times with *perceived opportunity*, 515 times with *tax avoidance*, 258 times with *tax perception*, etc. This adds up to a total connectivity score of 4097 for the concept *tax evasion*. Hence, the sum of the connectivity scores for the *Tax Evasion* theme = tax evasion 4097 + tax rate 3266 + probability of detection 2807 + trust 1646 + tax inducement 1982 + tax fairness 1745 + power 1107 + tax perception 1677 + tax enforcement 995 + income level 709 + reciprocity 108 + peer influence 127 = 20266. Count and connectivity provides different information. For example, a concept can appear very often (high count) but only in a very particular context, so its hit count would be high but its connectivity would be low. The connectivity is important as it shows what variables are being researched simultaneously.

Figure 3: Leximancer Interface showing the Connectivity for the Concept Tax Evasion

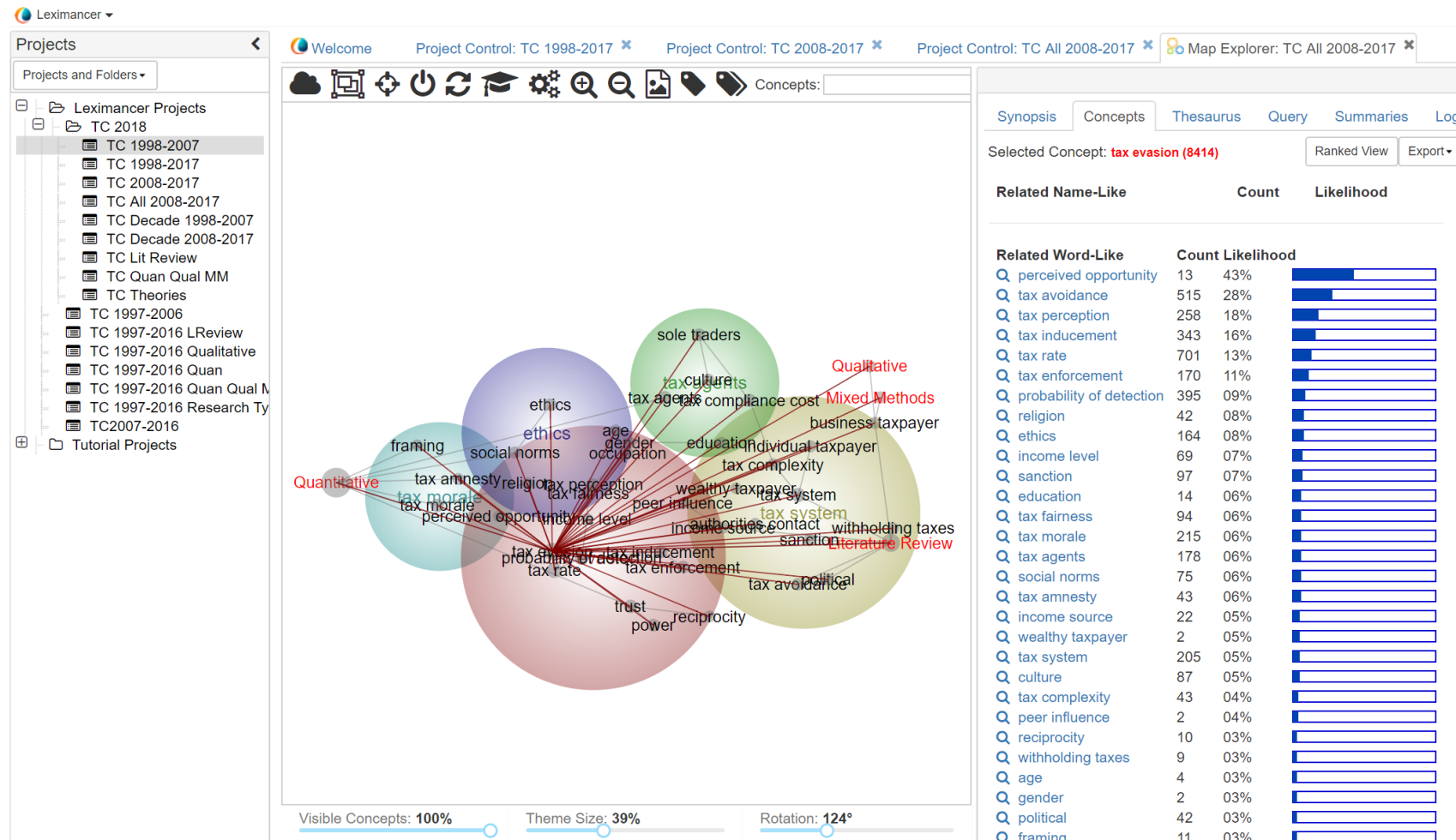


Table 9 demonstrates that *tax evasion*, *tax rate*, *tax system*, *probability of detection* and *tax morale* are the top five concepts across the entire data set. Further exploration of the indirect relationships between the concepts and their connectivity is detailed in Table 11.

Table 11: Top 5 Concepts and Their Connected Concepts



Top 5 Concepts	Total Count	Connected Concepts	Co-occurrence with Connected Concepts	Total Count of Connected Concepts	Connectivity Percentage
Tax evasion	8414	Perceived opportunity	13	30	43%
		Tax avoidance	515	1814	28%
		Tax perception	258	1414	18%
		Tax inducement	343	2197	16%
Tax system	4363	Tax fairness	361	1540	23%
		Tax complexity	239	1042	23%
		Wealthy taxpayer	7	42	17%
Probability of detection	4292	Peer influence	19	56	34%
		Authorities contact	18	65	28%
		Tax inducement	533	2197	24%
Tax rate	3884	Authorities contact	15	65	23%
		Peer influence	12	56	21%
		Tax inducement	278	2198	13%
Tax Morale	3560	Religion	91	516	18%
		Culture	170	1868	09%
		Trust	170	3465	05%

Tax evasion

Tax evasion (count = 8414) has a strong association with other identified concepts such as *perceived opportunity* (count = 30). These two concepts were coded 13 times together within the same sentence block across the data set, with a connectivity score of 43%. This score is calculated based on their co-occurrence score (i.e. 13) divided by the total number of times the connected concept (i.e. *perceived opportunity*) is coded across the entire data set (i.e. 30). However, it is important to note that a query search for the concepts ‘tax evasion’ and ‘perceived opportunity’ shows that of the 13 co-occurrences for these two concepts, 11 co-occurrences were from the same source - a

quantitative survey research article by Keraro et al..⁷⁰ This could suggest that while Keraro et al. considered these two concepts closely, other work largely did not. Keraro et al.'s study considered what effect, if any, perceived opportunity had for tax evasion on tax compliance for small and medium enterprises (SME) in Kenya.⁷¹ What the study concluded was that as the perceived opportunity for tax evasion increases, tax compliance declines among SMEs.

Tax evasion is also strongly connected with *tax avoidance* (count = 1814), with a co-occurrence score of 515 and connectivity score of 28%. It also has a strong connection with *tax perception* (count = 1414), with a co-occurrence score of 258 and connectivity score of 18%. The following statements illustrate how the concept *tax evasion* is used with these two other concepts of *tax avoidance* and *tax perception*:

“The main difference between tax evasion and tax avoidance is not only the interpretation of the laws but the exploitation of legal ambiguities by powerful taxpayers.”⁷²

“The dividing line between tax avoidance (legitimate tax planning) and tax evasion (abuse) it is not a natural or unanimous one, but rather depends on policy decisions of “each state.”⁷³

“If the perception becomes widespread that the government is not willing to detect and penalize evaders, then such a perception legitimizes tax evasion.”⁷⁴

“In contrast, equity theory predicts that a given level of tax payments, as the increase in public goods improves the taxpayer's perception of exchange equity, the taxpayer will be less likely to participate in tax evasion.”⁷⁵

Further exploration of the concept *tax evasion* using a query in Leximancer produced the following statements from the text that contributes to *tax evasion*:

“Based on Jackson and Hatfield (2005) and Highhouse and Yuce (1996), the risky alternative of tax evasion (or under-reporting) will be perceived as an

⁷⁰ J Keraro, M Oloo and P Ragama, “Effect of Perceived Opportunity for Tax Evasion on Tax Compliance among Small and Medium Enterprises in Nakuru Central Business District, Kenya” (2017) 2(1) *Mara Research Journal of Business & Management* 80.


⁷¹ Keraro, Oloo and Ragama, above n 70.

⁷² M Bergman, “Criminal Law and Tax Compliance in Argentina: Testing the Limits of Deterrence” (1998) 26(1) *International Journal of the Sociology of Law* 55, at 57.

⁷³ P.A. Barreto and C.A. Takano, “The Prevention of Tax Treaty Abuse in the BEPS Action 6: A Brazilian Perspective” (2015) 43(12) *Intertax* 825, at 829.

⁷⁴ J Alm and M McKee, “Extending the Lessons of Laboratory Experiments on Tax Compliance to Managerial and Decision Economics” (1998) 19(4/5) *Managerial & Decision Economics* 259, at 265.

⁷⁵ C.K. Kim, “Does fairness matter in tax reporting behavior?” (2002) 23(6) *Journal of Economic Psychology* 771, at 772.

"opportunity" by taxpayers in the loss domain and as a threat to taxpayers in the gain domain."⁷⁶ 

"Tax officers believed that taxpayers' intention, role of tax agents, taxpayers' ignorance of tax law and financial strain were the main reasons for non-compliance."⁷⁷

"This research seeks to incorporate into the compliance analysis psychological factors such as shame, embarrassment, and guilt, as well as taxpayer perceptions of fairness, satisfaction with the Government, and moral commitments to honest reporting."⁷⁸

The following statements demonstrate the tension between the two terms:

"Tax non-compliance can be divided into tax avoidance and tax evasion."⁷⁹

"Tax evasion and tax avoidance are deviations from what is to be found legal, fair and just."⁸⁰

"Tax evasion and tax avoidance concern more than simply the problem of losing public revenues."⁸¹

Tax system

The second most important concept *tax system* (count = 4363), is also closely connected to *tax fairness* (count = 1540, co-occurrence = 361, connectivity = 23%), *tax complexity* (count = 1042, co-occurrence = 239, connectivity = 23%) and *wealthy taxpayer* (count = 42, co-occurrence = 7, connectivity = 17%). The following statements illustrate this relationship:

"Tax scholars, at least those that follow in the footsteps of Kenneth Carter, often manifest a preoccupation with tax equity and fairness as a preliminary step on the way to advocating the importance of a robust, comprehensive income tax system."⁸²

⁷⁶ D.L. Sanders, P.M.J. Reckers and G.S. Iyer, "Influence of Accountability and Penalty Awareness on Tax Compliance" (2008) 30(2) *Journal of the American Taxation Association* 1, at 8.

⁷⁷ H Abdul-Jabbar and J Pope, "Exploring the Relationship Between Tax Compliance Costs and Compliance Issues in Malaysia" (2008) 1 *Journal of Applied Law and Policy* 1, at 5.

⁷⁸ S.W. Mazza, "Taxpayer Privacy and Tax Compliance Symposium: Tax Compliance: Should Congress Reform the 1998 Reform Act" (2002) 51(5) *University of Kansas Law Review* 1065, at 1068.

⁷⁹ M Suffian, T.M. Shamsudin, S.M. Sanusi, Z.M. and A.A. Hermawan, "Malaysian Code of Corporate Governance and Tax Compliance: Evidence from Malaysia" (2017) 16(2) *Management & Accounting Review* 157, at 159.

⁸⁰ D DeÁK, "Legal Considerations of Tax Evasion and Tax Avoidance" (2004) 26(1) *Society and Economy* 41, at 41.

⁸¹ DeÁK, above n 80, at 64.

⁸² K Brooks, "A Retrospective on the Contributions of Neil Brooks: So Far" (2015) 52 *Osgoode Hall Law Journal* 379, at 383.

“The taxpayer is thus left in the predicament where in order to be compliant in a system of increasing complexity, he or she must pay tax specialists to interpret and manage tax affairs.”⁸³

“When taxpayers believe the tax system to be unfair, the most common reasons are: the existence of too many loopholes; the wealthy and large corporations can evade taxes by hiring accountants and using tax shelters; and, the exchange of tax dollars for public goods and services is unequal.”⁸⁴

Probability of detection (also referred to as tax audits)

The third most important concept *probability of detection* (count = 4292) is closely connected to *peer influence* (count = 56, co-occurrence = 19, connectivity = 34%), *authorities contact* (count = 65, co-occurrence = 18, connectivity = 28%) and *tax inducement* (also defined as tax penalties) (count = 2197, co-occurrence = 533, connectivity = 24%). This illustrated in the following statements:

“However, beliefs about the probability of apprehension may also be affected by personal experience, peer influence and mass media exposure.”⁸⁵

“The last dimension is the tax system or structure which deals with complexity of the tax system, contact with tax authorities, sanctions, detection probability and tax rates.”⁸⁶

“Allingham and Sandmo’s (1972) classical model of tax evasion assumes that behaviour is influenced by factors such as the tax rate (which determines the benefits of evasion), the penalties for fraud, as well as the probability of detection which determine the cost).”⁸⁷

⁸³ J.A. Langham, N Paulsen and C.E.J. Härtel, “Improving tax compliance strategies: Can the theory of planned behaviour predict business compliance?” (2012) 10(2) *eJournal of Tax Research* 36, at 393.

⁸⁴ S King and S.M. Sheffrin, “Tax Evasion and Equity Theory: An Investigative Approach” (2002) 9(4) *International Tax and Public Finance* 505, at 512.

⁸⁵ K Devos, “Measuring and Analysing Deterrence in Taxpayer Compliance Research” (2007) 10(2) *Journal of Australian Taxation* 182, at 205.

⁸⁶ D Ho and B Wong, “Issues on compliance and ethics in taxation: what do we know?” (2008) 15(4) *Journal of Financial Crime* 369, at 372.

⁸⁷ R Oberholzer and E.M. Stack, “Perceptions of taxation: A comparative study of different population groups in South Africa” (2014) 40(2) *Public Relations Review* 226, at 227.

Tax rate

According to the Leximancer ranked concept list, the fourth most important concept was *tax rate* (count = 3884).⁸⁸ The concept *tax rate* (count = 3884) is closely connected to *authorities contact* (count = 65, co-occurrence = 15, connectivity = 23%), *peer influence* (count = 56, co-occurrence = 12, connectivity = 21%) and *tax inducement* (count = 2197, co-occurrence = 278, connectivity = 13%). The following statements illustrate this relationship:

“Specifically, an increase in the probability of tax detection is associated with a decrease in the tax rate.”⁸⁹

“This is due to an awareness that the analysis research using only deterrence variables (such as tax audits, tax rates, and tax penalties) are not sufficiently able to explain the level of tax compliance.”⁹⁰

Tax Morale

The fifth important concept *tax morale* (count = 3560) is closely connected to *religion* (also defined as religiosity, count = 51, co-occurrence = 91, connectivity = 18%), *culture* (count = 1868, co-occurrence = 170, connectivity = 9%), and *trust* (count = 3465, co-occurrence = 170, connectivity = 5%). This is illustrated in the following statements:

“Most previous studies have used available data including from the *World Values Survey* to study the role of religiosity in tax compliance or, more specifically, tax morale.”⁹¹

“However, more recent research shows that former East Germans demonstrate higher tax morale than former West Germans and that “the cultural background seems to have an effect on tax morale”.⁹²

⁸⁸ Initially the analysis demonstrated that for the concept ‘tax rate’ in the coded text blocks shows that the term ‘rate’ was used also to denote ‘audit rate’ (355 times), ‘penalty rate’ (199 times), ‘compliance rate’ (203 times), ‘evasion rate’ (51 times), and ‘unemployment rate’ (24 times). We did not isolate coding of other terms for this compound concept as they were also important to the study. The query results show that the term ‘tax rate’ appeared only 3884 times in the coded text blocks, which decreases the ranking for this concept to four instead of two.

⁸⁹ F.S. Hung, “Tax Evasion, Financial Dualism and Economic Growth” (2015) 43(2) *Academia Economic Papers* 175, at 206.

⁹⁰ I.N. Darmayasa and Y.R. Aneswari, “The Ethical Practice of Tax Consultant Based on Local Culture” (2015) 211 *Procedia - Social and Behavioral Sciences* 142, at 143.

⁹¹ J Pope and R Mohdali, “The role of religiosity in tax morale and tax compliance” (2010) 25(4) *Australian Tax Forum* 565, at 585.

“Torgler (2003) found evidence that trust in government, pride, and religious conviction have systematic positive influences on tax morale even after controlling for a range of individual characteristics.”⁹³

Comparison with Jackson & Milliron and Richardson & Sawyer

To further understand the changes in our understanding of the tax compliance concepts between the two decades, a comparison is also made with Jackson & Milliron and Richardson & Sawyer in Table 12.

Table 12 reveals an overall orientation toward *tax evasion*, *tax rate*, *tax system*, *tax morale*, and *probability of detection* as determinants of tax compliance. There was an early emphasis on *tax agents*, *tax compliance cost*, *ethics*, *tax inducement*, *sanction*, *tax fairness* and *political* in 1998 - 2007. In 2008 - 2017, there has been a strong emphasis on *trust*, *power*, *sole traders and culture*. The concepts of *tax enforcement*, *religion*, *individual taxpayer*, *reciprocity* and *withholding taxes* also have slightly increased in prominence.

Jackson and Milliron identified fourteen main compliance variables: two of these variables have continued to receive significant attention in the period covered by this study (*probability of detection* and *tax rate*). A further four have remained prominent (*tax fairness*, *tax complexity*, *ethics* and *sanctions*), while others have not (*age*, *gender*, *education*, *income level*, *income source*, *occupation*, *peer influence* and *authority contact*). Richardson and Sawyer added five other variables: three have remain prominent since 1998 (*tax compliance cost*, *tax agents* and *tax inducements*), but not *tax amnesty* and *framing (of tax decisions)*. This finding concurs with the suggestion by Jackson and Milliron that the first four variables of ‘tax compliance cost, tax agents and tax inducements and tax amnesty’ were potential growth areas in the tax compliance literature.⁹⁴

⁹² S Yong and F Martin, “Tax compliance and cultural values: the impact of ‘individualism and collectivism’ on the behaviour of New Zealand small business owners” (2016) 31(2) *Australian Tax Forum* 289, at 297.

⁹³ N Hashimzade, G.D. Myles and B Tran-Nam, “Applications of Behavioural Economics to Tax Evasions” (2013) 27(5) *Journal of Economic Surveys* 941, at 972.

⁹⁴ B.R. Jackson and V.C. Milliron, “Tax compliance research: findings, problems and prospects” (1986) 5 *Journal of Accounting Literature* 125.

Table 12: Comparison of Tax Compliance Variables Coded in the 1998 – 2007 Decade and the 2008 – 2017 Decade

		1 st Decade		2 nd Decade		Total Period	
		1998-2007		2008-2017		1998-2017	
Study	Concepts	n	%	N	%	n	%
Jackson and Milliron (1986)	1. age	67	2%	71	1%	138	2%
	2. gender	40	1%	32	1%	72	1%
	3. education	101	3%	116	2%	217	3%
	4. income level	423	14%	592	11%	1015	12%
	5. income source	195	7%	249	5%	444	5%
	6. occupation (or status)	45	2%	93	2%	138	2%
	7. peer influence	27	1%	29	1%	56	1%
	8. ethics	1051	36%	1021	19%	2072	25%
	9. tax fairness	647	22%	893	16%	1540	18%
	10. tax complexity	334	11%	708	13%	1042	12%
	11. authorities contact	49	2%	16	0%	65	1%
	12. sanction	730	25%	718	13%	1448	17%
	13. probability of detection (or tax audit)	1646	56%	2646	48%	4292	51%
	14. tax rate	2269	77%	3207	59%	5476	65%
Richardson and Sawyer (2001)	1. tax compliance cost	1076	36%	1730	32%	2806	33%
	2. tax agents	1321	45%	1678	31%	2999	36%
	3. framing (of tax decisions)	210	7%	222	4%	432	5%
	4. tax inducement (or tax penalties)	945	32%	1252	23%	2197	26%
	5. tax amnesty	273	9%	496	9%	769	9%
Additional Tax Compliance Variables identified in this study	1. tax evasion	2950	100%	5464	100%	8414	100%
	2. tax system	1619	55%	2744	50%	4363	52%
	3. tax morale	1353	46%	2207	40%	3560	42%
	4. tax avoidance	670	23%	1144	21%	1814	22%
	5. tax enforcement ↑	478	16%	1132	21%	1610	19%
	6. power ↑↑	351	12%	1765	32%	2116	25%
	7. trust ↑↑	586	20%	2879	53%	3465	41%
	8. political	654	22%	902	17%	1556	18%
	9. social norms	523	18%	746	14%	1269	15%
	10. religion (or religiosity)↑	116	4%	400	7%	516	6%
	11. culture ↑	541	18%	1327	24%	1868	22%
	12. tax perception	500	17%	914	17%	1414	17%
	13. perceived opportunity	3	0%	27	0%	30	0%
	14. sole traders ↑↑	505	17%	1934	35%	2439	29%
	15. individual taxpayer↑	286	10%	746	14%	1032	12%
	16. business taxpayer	317	11%	446	8%	763	9%
	17. wealthy taxpayer	26	1%	16	0%	42	0%
	18. reciprocity ↑	90	3%	238	4%	328	4%
	19. withholding taxes↑	101	3%	200	4%	301	4%
Total		23118		41000		64118	

Note: (a) n = the number of times a concept is coded within the time period; (b) % = n divided by the number of times the top concept is coded for that decade; and (c) ↑↑ = Major increase in prominence, ↑ = Some increase in prominence

Our study has identified another 19 variables which appear in studies from 1998 (*1. tax evasion, 2. tax system, 3. tax morale, 4. tax avoidance, 5. tax enforcement, 6. power, 7. trust, 8. political, 9. social norms, 10. religion / religiosity, 11. culture, 12. tax perception, 13. perceived opportunity, 14. sole traders, 15. individual taxpayer, 16. business taxpayer, 17. wealthy taxpayer, 18. Reciprocity and 19. withholding taxes*). The tax compliance variables that have remained prominent or received greater attention over the 1998 – 2017 period are:

- *Tax evasion, tax system, tax morale, power and trust* have received major attention in the tax compliance literature since 1998,
- *Sole traders* and *culture* have increased in prominence in 2008 – 2017 with an increase in likelihood score from 17% to 35%, and 18% to 24%, respectively,
- *Individual taxpayers* and *business taxpayers* remained prominent in 2008 – 2017 (with likelihood scores of 14% and 8%, respectively),
- Although *religion, reciprocity and withholding taxes* concepts have low occurrence scores when compared to other concepts, these concepts have increased in prominence over the last 20 years (with an increase in relevance scores from 4% to 7%, 3% to 4%, and 3% to 4%, respectively)

Overall, although some concepts have lost prominence over the years (*e.g., age, gender, occupation, education, income level, income source, peer influence, perceived opportunity*) further examination revealed that these concepts are relevant through indirect relationships they have with other more prominent concepts. For instance:

Age, gender, occupation and religion/religiosity are connected to tax morale:

“More recently tax morale has also been related to the characteristics of the environment in which the taxpayer takes his decision and a substantial effort has been exerted in order to identify external variables that play a significant role in shaping tax morale besides that played by individual characteristics (age, religiosity, gender, marital and occupational status) that have been regularly found to.”⁹⁵

⁹⁵ A Filippin, C.V. Fiorio and E Viviano, “The effect of tax enforcement on tax morale” (2013) 32 *European Journal of Political Economy* 320, at 320.

Culture and religion are also connected to *tax morale*:

“The tax morale part of the diagram at Figure 2 is now extended (upwards) to include six main attitudinal factors, including norms (encompassing what some researchers prefer to call “culture”), religiosity, education, attitudes to government/tax authority and to society, and perception of fairness.”⁹⁶

Peer influence is grouped together with *tax fairness*:

“Based on the fourteen tax compliance variables identified by Jackson and Milliron (1986), Fisher, Wartick and Mark (1992) have grouped ethics or morality, tax fairness and peer influence together representing factors influencing an individual’s tax compliance behaviour.”⁹⁷

Section Five: Limitations, Recommendations and Conclusion

Limitations of this research and opportunities for future research

With all studies there are limitations in the analysis and scope of the research as well as areas for future research. This applies even in instances where the study comprises a comprehensive literature review.


Limitations

Leximancer analysed the entirety of each article (including the literature review that many journal articles will have); this means part of the analysis presented is not just the results of the ‘new’ research, but also the summary of the prior research. It would be interesting to see whether analysis of only the ‘method’, ‘results’ and ‘conclusion’ sections of articles would produce different results. This could more clearly present how our knowledge has developed over time. The fact that prior research is included in the analysis may significantly distort the outcomes of this research. There is a strong argument in favour of taking the approach of first conducting some sampling to see whether this is a significant limitation. Such tests for robustness could be used in future iterations of this project.

⁹⁶ J Pope and M McKerchar, “Understanding tax morale and its effect on individual taxpayer compliance” (2011) 5 *British Tax Review* 587, at 597.

⁹⁷ Pope and Mohdali, above n 91, at 567.

Another limitation of Leximancer is that the researcher still needs to go to source documentation to ascertain the directional effect (if any) of the concepts discussed and the topic of tax compliance. Additionally, one must be mindful of jurisdictional differences that could influence whether the results would be replicable in other jurisdictions. Additionally, articles (especially as they are sourced from a variety of disciplines), may use inconsistent terms or attach different meanings to concepts they discuss.

In a similar manner to Richardson & Sawyer,⁹⁸ no books, discussion papers or conference papers form part of the data analysed. Furthermore, no reports and working papers, and the like, from international organisations such as the OECD, IMF, World Bank, revenue authorities, have been included. As noted earlier, works from these organisations have been significant contributors to the broader literature. Also, only  articles published in English were analysed, along with only peer-reviewed articles with full text available from internationally respected databases. Through these restrictions imposed on the analysis it is expected that some important contributions may have been excluded. In some instances these could form part of a future review if the works are subsequently published in peer reviewed journals. In other cases, the absence of these contributions will be a limitation on the analysis. The substantial increase in published research on tax compliance has also meant the level of analysis reproduced in this article is significantly less than that of Richardson and Sawyer.⁹⁹ That said, the other insights that Leximancer provides are, we believe, valuable.

A further limitation is that this article examines only the tax compliance variables. Due to the immense growth in the literature, analysis of the methodological approaches, along with the theories adopted, is beyond the scope of the analysis here. This is to be addressed in future work. We accept that theory development has always been ‘weak’ in the field of taxation compliance research. It is interesting to ask why this has been the case? Is it because tax is a recent academic discipline?¹⁰⁰ Are there any signs of change? For example, why has tax compliance not capitalised on the

⁹⁸ Richardson and Sawyer, above n 7.

⁹⁹ Richardson and Sawyer, above n 7.

¹⁰⁰ S Hayes, B Freudenberg and D Delaney, “Role of Tax Knowledge and Skills: What are the graduate skills required by small to medium Accounting firms” (2018) 13(1) *Journal of Australasian Tax Teachers Association* 152, at 157-158.

theoretical developments in areas such as criminology. Criminology is a theoretically rich new discipline that has moved past the stage of descriptive, empirical work. As a consequence of the limitation in scope of this article to tax compliance variables, the article has in some respects limited capacity to identify theory development in the field. While this at first instance may seem to be unfortunate and should have been avoided, the sheer size of the tax compliance literature renders inclusion of this aspect beyond the scope of a single article. It is our intention to address this in future research, as well as shed light on the extent to which the tax compliance research field operates in a silo with respect to other fields of scholarly endeavour.

Future research

With the substantial growth in the number of published studies, it is vital to make use of technology that assists in analysing the research. In this regard use of tools such as Leximancer are essential to not only reduce the sheer number of hours of analysis, but also to ensure the major themes are captured. We would also encourage future researchers to use Leximancer to assist with undertaking comparative analysis with our study.

Future research could consider the development of the methodologies utilised over the study period, and compare these to those identified by Richardson and Sawyer. Additionally, future research could consider which theories have underpinned the tax compliance studies from 1988 to 2017. Such analysis will be a future focus of the authors.

Future research could also analyse all of the concepts/factors and associated variables in detail, and undertake this within two ten year blocks. Additionally, studies reported in languages other than English could be analysed. Other databases, beyond the present 15 used in this study, could be explored for publications not captured in this study. Additionally, tax compliance studies published in books, as well as conference papers and other respected outlets, could also be analysed.

Other taxes

An analysis of compliance studies about fuel tax, inheritance tax and land tax could also be undertaken. The last types of taxes could be of particular importance given that we are witnessing potentially the largest inter-generational wealth transfer, as well as

arguments for an increased emphasis on land tax given concerns about the mobility of capital, and what this means for the income tax base.

Recommendations

Undertakings of this magnitude are not expected to be regular events. The growth in studies reviewed from 43 by Jackson & Milliron,¹⁰¹ to approximately 170 by Richardson & Sawyer,¹⁰² are relatively small in comparison to this study (713) as at the end of 2017. That said, part of the reason for the increase is that this study considers 20 years compared to ten years by Richardson & Sawyer. To make future reviews of the tax compliance literature less time consuming and more timely, updates should be taken more often, perhaps once each decade (especially as the average number of papers has increased each decade).

Given the results it appears that to understand tax compliance we need to better understand the counter-concept of ‘tax evasion’. The analysis reveals that the concept of tax evasion is highly connected to the concepts of tax rate, probability of detection and trust. It appears that if we can best understand the interaction of the concepts then we could better understand tax compliance overall. For example, Saad, in a study of New Zealand and Malaysian taxpayers, found different responses to two possible tax avoidance scenarios.¹⁰³ In a circumstance of ‘overstating business deductions’ there was less difference in the avoidance rates between taxpayers. This in part could be related to the possibility of detection given the likely document trail. In contrast, for ‘understating cash income’, there was a significant difference of compliance rates between those with higher perceptions of fairness of the tax system compared with taxpayers with lower perceptions. This in part may have been attributed to the perception of a low possibility of detection of ‘cash income’. Consequently, the probability of detection may play a large role amongst taxpayers as to whether to engage in tax evasion. In the event of low probability of detection, then the decision whether to comply or not could in part depend on perceptions of fairness.

¹⁰¹ Jackson and Milliron, above n 6.

¹⁰² Richardson and Sawyer, above n 7.

¹⁰³ N Saad, *Fairness Perceptions and Compliance Behaviour: Taxpayers’ Judgements in Self-Assessment Environments*. A thesis submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy in Taxation in the University of Canterbury, University of Canterbury (NZ), 2011.

With the formation of this list of top concepts, future researchers can reflect upon what factors and concepts have been researched in the past and then consider which ones should be explored in their own studies. Also, it allows revenue authorities to reflect on their own strategies to consider how they adjust them to greater enhance compliance. For example, with advancements in technology, there is greater data available for revenue authorities to know taxpayers' income (such as wages, interest and dividends), then in this way they can assist taxpayers to pre-fill tax returns which could raise awareness in taxpayers that their probability of detection is higher.

With the availability of powerful tools such as Leximancer, future researchers should consider expanding the source of research to include books and book chapters, as well as work published by major organisations (such as the OECD, World Bank and IMF), along with unpublished work such as working papers.

Additionally, this analysis may raise questions of what areas need further investigation. For example, the role of tax literacy and education does not have a major presence in the results. With current moves arguing for the need for greater tax literacy, the role this does or does not have on tax compliance should be explored.

Conclusions

Part of the importance of research is to add to our understanding of a phenomenon, namely tax compliance. In this regard it is important to take stock and analyse how our knowledge and understanding has evolved. Over 30 years ago Jackson & Milliron set about to review in the USA the understanding of tax compliance, and in their analysis of 43 paper they identified 14 variables. Then Richardson & Sawyer set about reviewing the decade after Jackson & Milliron from 1986 to 1997. By doing so, they reviewed nearly four times the number of papers from a wider jurisdictional basis and identified five additional variables.

This current study follows these two prior studies (in adopting a similar methodological approach), and analyses how our understanding of tax compliance has 'matured' over the two decades since Richardson and Sawyer's review. We acknowledge that there have been other significant contributions to the tax compliance literature in the interim, although these have adopted different approaches in their

analysis.¹⁰⁴ These studies are integral to our knowledge and understanding of tax compliance (and non-compliance).

The use in this study of Leximancer software meant the analysis was undertaken on a more ‘objective’ basis. In our study not only has there been an increase in the number of compliance variables examined, but the relative emphasis on variables examined by Jackson and Milliron, and Richardson and Sawyer, has altered. Through this analysis it was identified that in addition to the 19 variables already identified in the prior two studies, another 19 variables had been used in research concerning tax compliance. In particular, of these new variables, it appeared that tax evasion is the most frequently researched item for the last two decades, followed by trust, tax system, tax morale and tax avoidance.



While tax compliance is an ever-evolving concept, with such a comprehensive understanding we are much better placed to understand how future researchers could approach taxpayer compliance. Tax compliance is a multi-faceted phenomenon that requires consideration of many aspects from the political to social norms. Through the meta-analysis that is possible using Leximancer, we are able to gain further and deeper understanding of concepts studied and their relationship with each other.

That said, we need to ask ourselves whether this field is largely perseverating on the same ideas without progressing the conceptual work significantly. We acknowledge that the proportion of standard replication or application is probably always going to be higher than innovation no matter how innovative the field. This is in part due to the academic paradigm that generally requires research to build upon prior research and established theories; which is best illustrated by the need of an extensive literature review. However, is there the potential for a paradigm shift such as with the combining of theories and approaches from economics and psychology? Also, being somewhat provocative, are the standards for publication in the field of tax compliance set ‘too low’ or ‘too safe’ since essentially the same ideas and concepts are being discussed over and over again? When will the next thought provoking contribution emerge that significantly challenges our understanding? We sincerely hope this will not be far away.



¹⁰⁴ See, for example, footnotes 29-31 above.

